DOI:
10.1001/jamapsychiatry.2017.0284

Document Version
Peer reviewed version

DEVELOPMENT AND VALIDATION OF A CLINICALLY BASED RISK CALCULATOR FOR THE TRANSDIAGNOSTIC PREDICTION OF PSYCHOSIS

Paolo Fusar-Poli MD PhD[1,2], Grazia Rutigliano MD[1,3], Daniel Stahl PhD[1], Cathy Davies MSc[1], Ilaria Bonoldi MD PhD[1,2], Thomas Reilly[1], Philip McGuire MD PhD[1].

Affiliations

1. Institute of Psychiatry, Psychology & Neuroscience, King's College London, 16 De Crespigny Park, SE5 8AF London, United Kingdom;
2. OASIS service, South London and the Maudsley NHS Foundation Trust, 190 Kennington Lane, SE11 5DL London, United Kingdom;
3. Department of Clinical and Experimental Medicine, University of Pisa, via Roma 67, 56100 Pisa, Italy.

Abstract: 347
Text: 3000
Tables: 3
Figures: 2

Endnote reference file: Transdiagnostic_final_CD.enl

Correspondence to: Dr Paolo Fusar-Poli, PhD, Department of Psychosis Studies, 5th Floor, Institute of Psychiatry, Psychology & Neuroscience, PO63, 16 De Crespigny Park, SE5 8AF London, UK.

E-mail: paolo.fusar-poli@kcl.ac.uk

KEY POINTS

Question

What are the factors modulating risk enrichment in help-seeking subjects referred for clinical assessment on suspicion of psychosis risk?

Findings

This cohort study that included 710 subjects assessed for suspected psychosis indicated substantial six-year risk enrichment (15%) and provided a stratification model that is based on ethnicity and source of referral.

Meaning

Stratification of risk enrichment in subjects undergoing assessment for suspected psychosis risk may inform outreach campaigns, subsequent testing and optimise psychosis prediction.

ABSTRACT

Objective

To measure the proportion of individuals with a first episode of psychosis detected by At Risk Mental State (ARMS) services in secondary mental health services. To develop and externally validate a practical web-based individualised risk calculator tool for the transdiagnostic prediction of psychosis in secondary mental health care.

Design

Clinical register-based cohort study.

Setting

Subjects were drawn from electronic, real-world, real-time clinical records relating to 2008-2015 routine secondary mental health care in South London and Maudsley (SLaM) National Health Service (NHS) Foundation Trust.

Participants

All patients receiving a first index diagnosis of non-organic and non-psychotic mental disorder within SLaM NHS Trust in the period 1st January 2008 to 31st December 2015.

Main outcome measure

Risk of development of non-organic ICD-10 psychotic disorders.

Results

91199 patients receiving a first index diagnosis of non-organic and non-psychotic mental disorder within SLaM NHS Trust were included in the derivation (33820) or external validation (54716) datasets. The mean follow-up was 1588 days. The overall 6-year risk of psychosis in secondary mental health care was 3.02 (95% CI 2.88 to 3.15), which is higher than in the local general population. Compared with the ARMS designation, all of the ICD-10 diagnoses showed a lower risk of psychosis, with the exception of bipolar mood disorders (similar risk) and brief psychotic episodes (higher risk). The ARMS designation accounted only for a small proportion of transitions to psychosis (52/1001=5.19% in the derivation dataset), indicating the need for transdiagnostic prediction of psychosis in secondary mental health care. A prognostic risk stratification model based on preselected variables, including

index diagnosis, age, gender, age by gender, and ethnicity was developed and externally validated, showing good performance and potential clinical usefulness.

Conclusions

The online individualised risk calculator can be of clinical usefulness for the transdiagnostic prediction of psychosis in secondary mental health care. The risk calculator can help to identify those patients at risk of developing psychosis who require an ARMS assessment and specialised care. The use of this calculator may eventually facilitate the implementation of an individualised provision of preventative focused interventions and improve outcomes of first episode psychosis.

Study registration: researchregistry1487 (www.researchregistry.com), July 31st 2016. Data analysis began on September 1st 2016.

INTRODUCTION

Existing treatments for psychotic disorders have little impact on the course of illness under standard care.[1,2] Prevention and early intervention may be the only available clinical possibility to alter the course of psychosis.[3] Prevention of psychosis has been feasible since the introduction of the At Risk Mental State (ARMS) construct, two decades ago.[4] The ARMS has been validated internationally,[5-7] and it can reliably identify young individuals at specific enhanced risk for the development of psychotic disorders[8] - mostly schizophrenia spectrum disorders[9] - but not of non-psychotic disorders,[10,11] over the following two to three years.[12] Randomised controlled trials have shown that focused interventions, if offered to ARMS individuals, can effectively reduce the risk of future illness.[13,14] Owing to these unprecedented potentials, the ARMS has gained traction to the point that specialised assessment and treatment is recognised as a key component of secondary mental health services by National Institute for Health and Care Excellence (NICE) guidelines.[15]

However, the overall clinical impact of the ARMS on psychosis prevention in secondary mental health care, and the value of using the ARMS designation as compared to standard mental diagnoses (e.g. those defined by the International Classification of Disease, 10th edition, ICD-10), is not completely clear. For example, the majority of ARMS individuals would also meet criteria for a secondary diagnosis of comorbid mental disorder, mostly depression or anxiety.[16] As a result, some authors have claimed that the ARMS construct is not strictly *necessary*[17], and that psychosis could be predicted (and treated) within the existing diagnostic categories of common mental disorders.[18] Whether we can pragmatically predict psychosis outside the ARMS designation or not remains unclear, because no studies have ever addressed this issue. Such a gap of knowledge may have clinical implications for the provision of preventative intervention services and policy makers. In fact, some authors suggest that it would be better to detect and treat psychosis as it emerges from common mental disorders rather than promoting new ARMS services.[18] To further compound the issue, the overall burden of psychosis risk in secondary mental health care is mostly undetermined,

and it is not clear whether the ARMS designation is *sufficient* to pragmatically detect all individuals who will later develop a first episode of psychosis. Since ARMS services usually receive referrals on suspicion of psychosis risk, it is possible that not all individuals in secondary mental health care who will later develop a first episode of psychosis would eventually be detected by ARMS services.

The current study advances knowledge by investigating, for the first time, the proportion of first episode individuals detected by ARMS services in secondary mental health services as well as the transdiagnostic risk of developing psychotic disorders across any ICD-10 defined mental disorder. The primary aim was to develop and validate a clinically based, practical, individualised risk calculator tool to facilitate the transdiagnostic prediction of psychosis in secondary mental health care and increase the proportion of individuals at risk for psychosis detected by ARMS services, improving outcomes of first episode psychosis.

METHODS

Data source

Clinical register-based cohort selected thorough a Clinical Record Interactive Search (CRIS) tool[19] (see eMethods 1).

Study population

All individuals accessing SLaM services in the period 1st January 2008 to 31st December 2015, and who received a first index primary diagnosis of any non-organic and non-psychotic mental disorder, were initially considered eligible. We then excluded those who developed psychosis in the three months immediately following the first index diagnosis. Approval for the study was granted by the Oxfordshire Research Ethics Committee C. Because the data set was made up of deidentified data, informed consent was not required[19].

Study measures

The outcome (risk of developing any psychotic disorder), predictors (index diagnosis, age, gender, ethnicity, and age by gender interaction), and time to event were automatically extracted using CRIS.[19] Predictors were preselected on the basis of previous meta-analytical clinical knowledge, as recommended[20] (see eMethods 2 and eTable 1 for full details).

Statistical analysis

This clinical register-based cohort study was conducted according to the REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement[21] (see supplementary checklist).

Baseline clinical and sociodemographic characteristics of the sample (including missing data) were described by means and SDs for continuous variables, and absolute and relative frequencies for categorical variables. The overall cumulative risk of psychosis onset in SLaM patients was described with the Kaplan–Meier failure function (1 – survival)[22] and Greenwood 95% CIs,[23] and was qualitatively compared with the risk of psychosis in the local general population (mean predicted cases across SLaM boroughs, estimated with PsyMaptic (http://www.psymaptic.org/)).[24]

Model development and validation followed the guidelines of Royston et al.,[25] Steyerberg et al.[26] and the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD).[27]

Model development

We used Cox proportional hazards multivariable complete-case analyses to evaluate the effects of the preselected predictors (index diagnosis, age, gender, ethnicity, and age by gender interaction) on the development of non-organic ICD-10 psychotic disorders and time to development of psychosis, after checking the proportional hazards assumption.[28] Model development was not based on stepwise methods, which are not recommended,[25] but on a-

priori selection of predictors based on previous knowledge, as detailed in the eMethods 2. Continuous variables were not dichotomised.[25] Because of significant sociodemographic differences between the SLaM boroughs (from[29]: see Table 1 and Figures 2 and 3), we used nonrandom split-sample development and external validation,[27] with the Lambeth and Southwark cases in the derivation sample and all other cases in the validation sample. The model with all preselected predictors was first fitted to the derivation data to estimate the optimal regression coefficients. Performance diagnostics of individual predictor variables in the derivation dataset were explored with Harrell's C-index,[25] which is similar to the area under the receiver operating characteristic curve. We then generated individual prognostic scores, allowing a prognostic index (PI) for risk of psychosis onset to be developed in the derivation dataset.[30] As a supplementary analysis we fitted the model after excluding the ATPD cases.

External model validation

The regression coefficients as estimated in the derivation dataset were then applied to each case in the external validation dataset, to generate the PI in the validation dataset. Overall model performance (the distance between the predicted outcome and actual outcome[26]) was assessed with the Brier score (the average mean squared difference between predicted probabilities and actual outcomes, which also captures calibration and discrimination aspects[26]). A lower score indicates higher precision and less bias, but interpretation depends on the incidence of the outcome.[26] Overall performance was further investigated with Royston's modification of Nagelkerke's R^2 (indexing the proportion of variation explained by the model).[31] Calibration (the agreement between observed outcomes and predictions[26]) was assessed with the regression slope of PI[26] (which also captures discrimination and model fit),[25] with the regression intercept (calibration-in-the-large,[26] estimated as previously detailed[32]) and with the calibration plot (resampling model calibration with hare function[33]). Discrimination (accurate predictions discriminate between those with and those without the outcome[26]) was addressed with Harrell's C-index[25] and with the discrimination slope

(difference in mean of predictions between outcomes[26]). Recent studies indicate that unbiased and precise estimation of performance measures can be achieved with a minimum of 100 events in the external validation dataset.[34]

Potential clinical usefulness for psychosis prevention

We additionally explored the potential clinical utility of the risk calculator as recently recommended by Vickers and Steyerberg.[35] Performance measures do not tell us whether the risk calculator would do more good than harm if used in clinical practice.[35] Net benefit (for details see[35] and[36]) analyses tackle such limitations by including an exchange rate, a clinical judgment of the relative value of benefits (such as preventing psychosis in secondary mental health care) and harms (such as unnecessary treatment) associated with the predictive model (see details on the exchange rate on eMethods 3). Since definition of the exchange rates is subjective, we additionally plotted net benefit for a range of reasonable exchange rates in a decision curve analysis, as recommended.[35]

All analyses were conducted in STATA 13 and R 3.3.0.

RESULTS

Sociodemographic and clinical characteristics of the sample

Of 92227 patients receiving a first index diagnosis of non-organic and non-psychotic mental disorder within SLaM in the period 2008-2015, 91199 fulfilled the study inclusion criteria and were included in the derivation or validation datasets, as indicated in Figure 1.

Table 1 shows the baseline characteristics of the study population, as well as the derivation and validation datasets. As expected, there were significant sociodemographic differences across the derivation and validation datasets (Table 1), particularly with respect to ethnicity and index diagnosis. The mean follow-up was 1588 days (95% CI 1582-1595) with no differences between the derivation and validation datasets (derivation: mean 1589, 95% CI

1579-1599; validation: mean 1588, 95% CI 1580-1596). The overall risk of developing a psychotic disorder is presented in the eResults 1 and eFigure 1 and the baseline hazard function in the eFigure 2.

Model development

In the derivation dataset there were 1001 transitions to psychosis (52 of which were observed in the ARMS, 5.19%, eTable 2). The multivariable model significantly predicted psychosis onset (likelihood ratio chi-square test=1767.59, $p<0.001$). Age and male gender were significantly associated with an increased risk of psychosis (Table 2). Across males, risk of psychosis decreased with increasing age (Table 2). Relative to White ethnicity, Black, Asian, mixed, and other ethnicities were associated with an increased risk of developing psychosis (Table 2). Compared with the reference ARMS designation, all of the other ICD-10 mental disorders were associated with a lower risk of developing psychosis, with two exceptions (Table 2). Bipolar mood disorders and acute and transient psychotic disorders showed a comparable and higher risk of psychosis than the ARMS, respectively (Table 2). Post-hoc analyses showed that relative to the ARMS, bipolar mood disorders and acute and transient psychotic disorders had a higher risk of developing affective psychoses (HR= 4.628, 95%CI 1.655 - 12.941), and schizophrenia spectrum psychoses (HR=5.457, 95%CI 2.742 - 10.893), respectively. Supplementary analyses using the APS subgroup of the ARMS as a reference group confirmed the model, showing that the BLIPS subgroup was at higher risk of developing psychosis than the APS subgroup. Model diagnostics using the C-index are detailed in Table 2. The model showed very good overall apparent performance (good discrimination, C-index 0.800) and explained approximately 74% of the observed variation (Table 3). The model remained significant after removing the ATPD cases (eTable 3).

Model validation

In the external validation dataset there were 1010 transitions to psychosis (12 of which were observed in the ARMS, 1.19%, eTable 2), a value that greatly exceeds the minimum of 100

events required for robust external validation.[34] The model retained an overall good performance and was able to explain around 72% of the observed heterogeneity (Table 3). Model discrimination was fair to good, with a C-index of 0.791 (Table 3). The mean risk of psychosis in the validation dataset was lower than in the derivation dataset, but there were no major miscalibration issues (Table 3 and eFigure 3).

Potential clinical usefulness of the risk calculator

At the reference threshold for recommending focused interventions to prevent psychosis, the use of the model was associated with significant net benefits in both the derivation and validation datasets (Table 3). The decision curve estimated in the validation dataset (Figure 2) shows that compared to conducting no tests, testing on the basis of the risk calculator is associated with net benefits for a 1-50% range of threshold probability (risk of developing psychosis by five years, see eResults 2 for an example).

An online version of the risk calculator was built to facilitate numeric calculation of the predicted probability of conversion to psychosis in secondary mental health care (http://www.psychosis-risk.net).

DISCUSSION

91199 patients receiving a first index diagnosis of non-organic and non-psychotic mental disorder within SLaM were included in the study, either in the derivation (33820) or validation (54716) datasets, with a mean follow-up of 1588 days. The overall 6-year risk of psychosis in secondary mental health care was 3.02 (95%CI 2.88-3.15) and which is higher than the 6-year risk of psychosis in the local general population (0.62). Compared with the ARMS designation, all of the other ICD-10 diagnoses were associated with a lower risk of psychosis, with two exceptions. Bipolar mood disorders and acute and transient psychotic disorders showed a similar and higher risk of psychosis than the ARMS, respectively. The ARMS designation accounted only for a small proportion of transitions to psychosis,

indicating the need for transdiagnostic prediction of psychosis in secondary mental health care. The prognostic risk stratification model based on preselected clinically based variables (index diagnosis, age, gender, age by gender, and ethnicity) showed good prognostic accuracy in the derivation dataset. The risk calculator was externally validated, confirming good performance and potential clinical usefulness for the transdiagnostic prediction of psychosis in secondary mental health care.

This study has three significant clinical implications. First, we confirmed substantial psychosis risk enrichment in individuals accessing secondary mental health care. The 6-year risk of psychosis was fivefold higher than in the local general population (3.02/0.62=4.87) and in individuals accessing primary care,[37] highlighting a clear window of opportunity for the transdiagnostic prevention of psychosis within secondary mental health care.[38]

Second, we have shown that the ARMS designation, in particular its APS subgroup (footnotes to Table 2) is necessary to predict psychosis in individuals who have never experienced psychotic (e.g. BLIPS[39]) symptoms. The ICD-10 categories of substance use disorders, non-bipolar mood disorders, anxiety disorders, personality disorders, developmental disorders, childhood/adolescence onset disorders, physiological syndromes, and mental retardation showed a lower level of psychosis risk. Accordingly, the use of current ICD-10 categories of comorbid mental disorders, such as anxiety or depression, is unlikely to be of any clinical usefulness to predict psychosis. ICD-10 acute and transient psychotic disorders and the BLIPS subgroup of the ARMS were both associated with a very high risk of developing psychosis, but only in individuals with remitting symptoms at the time of the index diagnosis.[40,41] Similarly, bipolar mood disorders specifically predicted the onset of affective-like psychoses.

Third, we have clearly demonstrated that the ARMS designation, although necessary, is not sufficient to intercept the overall burden of psychosis risk in secondary mental health care

(see eDiscussion 1). In fact, although OASIS was established in Lambeth and Southwark several years before the start of the current cohort,[42] only 314 individuals (out of 33820, 0.93%) were under OASIS care, accounting for only 5.19% of the total cases of emerging psychosis across the two boroughs. More importantly, none of the patients outside of OASIS care had ever been assessed for an ARMS. This seems like a missed clinical opportunity, because screening for an ARMS is specifically indicated for individuals "already distressed by mental problems"[43] and accessing secondary mental health care,[44] to prevent psychosis with focused interventions,[13] and improve outcomes in those who go on to develop psychosis (by reducing the duration of untreated psychosis, admission to hospital and compulsory treatments[45] or unnecessary treatment[46]).

Building on the aforementioned points, our findings highlight a significant unmet need for transdiagnostic prevention of psychosis in secondary mental health care, which is not currently addressed by existing ICD-10 categories (that are not specific enough for psychosis prediction) or the ARMS designation (which does not include the majority of individuals at risk for psychosis). To overcome these limitations, this study developed a practical, individualised risk calculator tool for the transdiagnostic prediction of psychosis in secondary mental health care. A well-performing risk calculator was developed from easily collectable clinical and demographic predictor variables (age, gender, age by gender, ethnicity, index ICD-10/ARMS diagnosis). The overall validated model was robust and achieved good performance, which is in the range of values for established calculators currently in use for cancer and cardiovascular, neurological and endocrine diseases (see Table 3 in[5]). The risk calculator was implemented online and designed to generate a number representing the probability of transition to psychosis, given a particular profile of input variables. A key advantage of the risk calculator is that it inherently accommodates heterogeneity in profiles of risk factors among high-risk individuals.[47,48] At the same time, the risk calculator assumes that individuals have accessed secondary mental health care and that the predictor variables are coded as indicated in our methods (e.g. ICD-10 categories for the index diagnosis).

Therefore, the risk prediction tool would not be usable in primary care or the general population, or if other diagnostic criteria have been used (e.g. DSM).

This tool is therefore most useful to clinicians using the calculator for patients who have accessed secondary mental health services. The online calculator could also be easily integrated into electronic case registers such as CRIS, to facilitate the automatic and individualised prediction of psychosis. Critically, risk determinations should be communicated to patients by clinicians who can help patients understand the meaning of the risk estimates and provide commensurate treatment recommendations. The decision curve analysis presented in our study can further help clinicians to tailor individualised focused interventions, such as selecting patients to be referred to ARMS services. Focused interventions may include a detailed clinical assessment for psychosis risk (i.e. the ARMS assessment) combined with sequential testing,[49,50] close-in clinical monitoring for the emergence of psychosis, and psychological treatments currently recommended to prevent psychosis.

Future studies are needed to refine the focused interventions targeting the high-risk individuals detected by our risk calculator. It is also possible that not all high-risk individuals, even if properly referred and assessed, would eventually meet ARMS criteria. For example, research in high-risk individuals with an index diagnosis of bipolar disorders may help to refine the ARMS construct and its ability to predict the onset of affective psychoses.[51] Similarly, the effectiveness of preventative psychological treatments in individuals deemed at risk by our calculator, but not meeting ARMS criteria, should be further investigated. Limitations of the study have been addressed in the eDiscussion 2 section.

CONCLUSIONS

Individuals accessing secondary mental health services are at enhanced risk of developing psychosis compared to the local general population. The use of this novel individualised risk

calculator can be of clinical usefulness for the transdiagnostic prediction and prevention of psychosis in secondary mental health care.

Details of contributors: PFP, GR, PM gave substantial contribution to concept and design; PFP, GR, DS, TR, CD, IB contributed to acquisition, analysis, and interpretation of data; PFP drafted the manuscript; PFP, DS, GR, TR, PM, CD, IB contributed to critical revision of the manuscript for important intellectual content; PFP, DS, GR, CD conducted the statistical analysis; PFP, DS, PM, IB contributed to supervision; PFP, GR, TR, PM, CD contributed to administrative, technical, or material support; PFP, PM obtained funding. PFP and GR had full access to all of the data in the study and can take responsibility for the integrity of the data and the accuracy of the data analysis.

Financial support: This study was supported in part by a 2014 NARSAD Young Investigator Award to PFP. The funder had no influence on the study design, collection, analysis and interpretation of the data, writing of the report and in the decision to submit the article for publication. The researchers are independent from the funder.

REFERENCES:

1. Millan MJ, Andrieux A, Bartzokis G, et al. Altering the course of schizophrenia: progress and perspectives. *Nat Rev Drug Discov.* 2016;15(7):485-515.

2. Jaaskelainen E, Juola P, Hirvonen N, et al. A systematic review and meta-analysis of recovery in schizophrenia. *Schizophrenia bulletin.* 2013;39(6):1296-1306.

3. Fusar Poli P, McGorry P, Kane J. Improving outcomes of first episode psychosis. *World Psychiatry.* 2017:in press.

4. Yung AR, McGorry PD, McFarlane CA, Jackson HJ, Patton GC, Rakkar A. Monitoring and care of young people at incipient risk of psychosis. *Schizophrenia bulletin.* 1996;22(2):283-303.

5. Fusar-Poli P, Cappucciati M, Rutigliano G, et al. At risk or not at risk? A meta-analysis of the prognostic accuracy of psychometric interviews for psychosis prediction. *World Psychiatry.* 2015;14(3):322-332.

6. Braham A, Bannour AS, Ben Romdhane A, et al. Validation of the Arabic version of the Comprehensive Assessment of At Risk Mental States (CAARMS) in Tunisian adolescents and young adults. *Early Interv Psychiatry.* 2014;8(2):147-154.

7. Krebs MO, Magaud E, Willard D, et al. [Assessment of mental states at risk of psychotic transition: validation of the French version of the CAARMS]. *Encephale.* 2014;40(6):447-456.

8. Fusar-Poli P, Schultze-Lutter F. Predicting the onset of psychosis in patients at clinical high risk: practical guide to probabilistic prognostic reasoning. *Evidence-based mental health.* 2016;19(1):10-15.

9. Fusar-Poli P, Bechdolf A, Taylor MJ, et al. At risk for schizophrenic or affective psychoses? A meta-analysis of DSM/ICD diagnostic outcomes in individuals at high clinical risk. *Schizophrenia bulletin.* 2013;39:923-932.

10. Webb JR, Addington J, Perkins DO, et al. Specificity of Incident Diagnostic Outcomes in Patients at Clinical High Risk for Psychosis. *Schizophrenia bulletin.* 2015;41(5):1066-1075.

11. Fusar-Poli P, Rutigliano G, Stahl D, et al. Long-Term validity of the at risk mental state (ARMS) for predicting psychotic and non-psychotic mental disorders. *European Psychiatry.* 2017:in press.

12. Kempton M, Bonoldi I, Valmaggia L, McGuire P, Fusar-Poli P. Speed of psychosis progression in people at ultra high clinical risk: a complementary meta-analysis. *JAMA Psychiatry.* 2015;72(6):622-623.

13. Stafford MR, Jackson H, Mayo-Wilson E, Morrison AP, Kendall T. Early interventions to prevent psychosis: systematic review and meta-analysis. *Bmj.* 2013;346:f185.

14. Deas G, Kelly C, Hadjinicolaou AV, Holt C, Agius M, Zaman R. An update on: meta-analysis of medical and non-medicaltreatments of the prodromal phase of psychotic illness in at risk mental states. *Psychiatr Danub.* 2016;28(Suppl-1):31-38.

15. NICE. Psychosis and schizophrenia in children and young people: Recognition and management. http://www.nice.org.uk/Guidance/CG155 2013; Jan 2013. Accessed 17 Jan 2016.

16. Fusar-Poli P, Nelson B, Valmaggia L, Yung AR, McGuire PK. Comorbid depressive and anxiety disorders in 509 individuals with an at-risk mental state: impact on psychopathology and transition to psychosis. *Schizophrenia bulletin.* 2014;40(1):120-131.

17. Conrad AM, Lewin TJ, Sly KA, et al. Utility of risk-status for predicting psychosis and related outcomes: evaluation of a 10-year cohort of presenters to a specialised early psychosis community mental health service. *Psychiatry Res.* 2017;247:336-344.

18. van Os J, Murray RM. Can we identify and treat "schizophrenia light" to prevent true psychotic illness? *Bmj.* 2013;346:f304.

19. Stewart R, Soremekun M, Perera G, et al. The South London and Maudsley NHS Foundation Trust Biomedical Research Centre (SLAM BRC) case register: development and descriptive data. *BMC psychiatry.* 2009;9:51.

20. Studerus E, Ramyead A, Riecher-Rossler A. Prediction of transition to psychosis in patients with a clinical high risk for psychosis: a systematic review of methodology and reporting. *Psychological medicine.* 2017:1-16.

21. Benchimol EI, Smeeth L, Guttmann A, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. *PLoS medicine.* 2015;12(10):e1001885.

22. Kaplan E, Meier P. Nonparametric estimation from incomplete observations. *J Amer Statist Assn.* 1958;53(282):457-481.

23. Greenwood M. *The natural duration of cancer.* London1926.

24. Kirkbride JB, Jackson D, Perez J, et al. A population-level prediction tool for the incidence of first-episode psychosis: translational epidemiology based on cross-sectional data. *BMJ open.* 2013;3(2).

25. Royston P, Altman DG. External validation of a Cox prognostic model: principles and methods. *BMC Med Res Methodol.* 2013;13:33.

26. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology.* 2010;21(1):128-138.

27. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. *Annals of internal medicine.* 2015;162(1):55-63.

28. Grambsch P, Therneau T. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika.* 1994;81:515-526.

29. Perera G, Broadbent M, Callard F, et al. Cohort profile of the South London and Maudsley NHS Foundation Trust Biomedical Research Centre (SLaM BRC) Case Register: current status and recent enhancement of an Electronic Mental Health Record-derived data resource. *BMJ open.* 2016;6(3):e008721.

30. Machin D, Cheung Y, Parmar M. *Survival Analysis: A Practical Approach.* 2nd Edition ed. Chichester, UK.: John Wiley & Sons Ltd.; 2006.

31. Royston P. Explained variation for survival models. *The Stata Journal.* 2006;6(1):83-96.

32. Crowson CS, Atkinson EJ, Therneau TM. Assessing calibration of prognostic risk scores. *Statistical methods in medical research.* 2014.

33. Harrell F. Regression Modeling Strategies. 2017. Accessed 13-01-2017.

34. Collins GS, Ogundimu EO, Altman DG. Sample size considerations for the external validation of a multivariable prognostic model: a resampling study. *Statistics in medicine.* 2016;35(2):214-226.

35. Vickers AJ, Van Calster B, Steyerberg EW. Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. *Bmj.* 2016;352:i6.

36. Vickers AJ, Cronin AM, Elkin EB, Gonen M. Extensions to decision curve analysis, a novel method for evaluating diagnostic tests, prediction models and molecular markers. *BMC Med Inform Decis Mak.* 2008;8:53.

37. Hardoon S, Hayes JF, Blackburn R, et al. Recording of severe mental illness in United Kingdom primary care, 2000-2010. *PloS one.* 2013;8(12):e82365.

38. Francesconi M, Minichino A, Carrion RE, et al. Psychosis prediction in secondary mental health services. A broad, comprehensive approach to the "at risk mental state" syndrome. *Eur Psychiatry.* 2016;40:96-104.

39. Fusar-Poli P, Cappucciati M, De Micheli A, et al. Diagnostic and Prognostic Significance of Brief Limited Intermittent Psychotic Symptoms (BLIPS) in Individuals at Ultra High Risk. *Schizophrenia bulletin.* 2017; 43(1):48-56.

40. Fusar-Poli P, Cappucciati M, Bonoldi I, et al. Prognosis of brief psychotic episodes: a meta-analysis. *JAMA Psychiatry.* 2016;73(3):211-220.

41. Fusar-Poli P, Cappucciati M, Borgwardt S, et al. Heterogeneity of psychosis risk within individuals at clinical high risk: A meta-analytical stratification. *JAMA Psychiatry.* 2016;73(2):113-120.

42. Fusar-Poli P, Byrne M, Badger S, Valmaggia LR, McGuire PK. Outreach and support in south London (OASIS), 2001-2011: ten years of early diagnosis and treatment for young individuals at high clinical risk for psychosis. *Eur Psychiatry.* 2013;28(5):315-326.

43. Schultze-Lutter F, Michel C, Schmidt SJ, et al. EPA guidance on the early detection of clinical high risk states of psychoses. *Eur Psychiatry.* 2015;30(3):405-416.

44. Fusar-Poli P, Rutigliano G, Stahl D, et al. Deconstructing risk enrichment to optimize prediction of psychosis in individuals at clinical high risk. *JAMA Psychiatry.* 2016:in press.

45. Valmaggia LR, Byrne M, Day F, et al. Duration of untreated psychosis and need for admission in patients who engage with mental health services in the prodromal phase.

The British journal of psychiatry : the journal of mental science. 2015;207(2):130-134.

46. Woodberry KA, Seidman LJ, Bryant C, et al. Treatment Precedes Positive Symptoms in North American Adolescent and Young Adult Clinical High Risk Cohort. *J Clin Child Adolesc Psychol.* 2016:1-10.

47. Carrion RE, Correll CU, Auther AM, Cornblatt BA. A Severity-Based Clinical Staging Model for the Psychosis Prodrome: Longitudinal Findings From the New York Recognition and Prevention Program. *Schizophr Bull.* 2017;43(1):64-74.

48. Addington J, Liu L, Perkins DO, Carrion RE, Keefe RS, Woods SW. The Role of Cognition and Social Functioning as Predictors in the Transition to Psychosis for Youth With Attenuated Psychotic Symptoms. *Schizophr Bull.* 2017;43(1):57-63.

49. Schmidt A, Cappucciati M, Radua J, et al. Improving prognostic accuracy in subjects at clinical high risk for psychosis: systematic review of predictive models and meta-analytical sequential testing simulation *Schizophr Bull.* 2016;Aug 17:Epub.

50. Clark SR, Baune BT, Schubert KO, et al. Prediction of transition from ultra-high risk to first-episode psychosis using a probabilistic model combining history, clinical assessment and fatty-acid biomarkers. *Transl Psychiatry.* 2016;6(9):e897.

51. Scott J, Marwaha S, Ratheesh A, et al. Bipolar At-risk Criteria: An Examination of Which Clinical Features Have Optimal Utility for Identifying Youth at Risk of Early Transition From Depression to Bipolar Disorders. *Schizophr Bull.* 2016.

Figure 1. Flow chart of study population

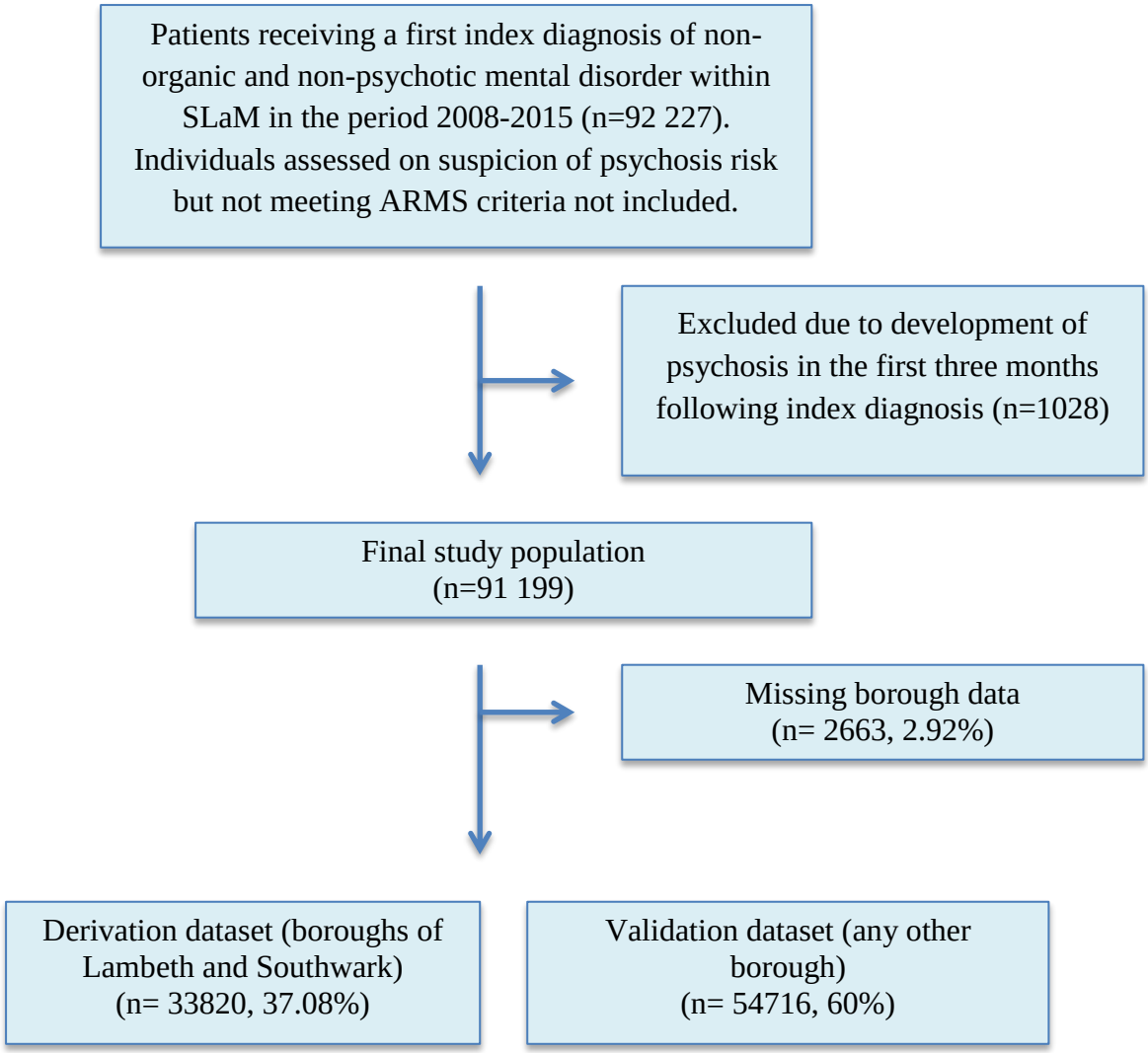

Table 1. Sociodemographic characteristics of study population, including the derivation and validation dataset

		Study population (n= 91199)(a)		Derivation dataset (n=33820)		Validation dataset (n=54716)		Validation vs Derivation	
		Mean	SD	Mean	SD	Mean	SD	t	P
Age (years)(a)		32.97	18.63	34.4	18.92	31.98	18.54	18.73	<0.001
		Count	%	Count	%	Count	%	X^2	P
Gender								13.37	<0.001
	Male	46404	50.88	17303	48.81	27302	49.9		
	Female	44761	49.08	16507	51.16	27398	50.07		
	Missing	34	0.04	10	0.03	16	0.03		
Ethnicity								50.21	<0.001
	Black	14327	15.71	6879	20.34	7023	12.84		
	White	55679	61.05	18627	55.08	35392	64.68		
	Asian	3830	4.2	1129	3.34	2608	4.77		
	Mixed	3319	3.64	1306	3.86	1957	3.58		
	Other	5700	6.25	3466	10.25	2084	3.81		
	Missing	8344	9.15	2413	7.13	5652	10.33		
Index diagnosis								48.2	<0.001
	ARMS	368	0.4	314	0.93	50	0.09		
	Acute and transient psychotic disorders	1370	1.5	553	1.64	725	1.33		
	Substance use disorders	14689	16.11	7149	21.14	6507	11.89		
	Bipolar mood disorders	2558	2.8	950	2.81	1526	2.79		
	Non bipolar mood disorders	15496	16.99	6302	18.63	8841	16.16		
	Anxiety disorders	24770	27.16	8235	24.35	15960	29.17		
	Personality disorders	3562	3.91	1286	3.8	2116	3.87		
	Developmental disorders	5192	5.69	1412	4.18	3706	6.77		
	Childhood/adolescence onset disorders	13984	15.33	4200	12.42	9629	17.6		
	Physiological syndromes	7053	7.73	2555	7.55	4424	8.09		
	Mental retardation	2157	2.37	864	2.55	1232	2.25		

(a) SLaM boroughs used to define the derivation (Lambeth and Southwark) and validation (any other) datasets: Lambeth and Southwark 33820 (37.08%), any others 54716 (60.00%), missing 2663 (2.92%)

Table 2. Statistics for individual predictor variables in the multivariable Cox proportional hazards regression analysis of risk for psychosis in the derivation dataset

Predictor	Multivariable model				Harrell's C (a)	
	Hazard Ratio	*95% CI*		*P*	*Decrement if removed*	*Increase if added (b)*
Age (years)	1.011	1.001	1.017	<0.001	<0.001	n/a
Gender					<0.001	0.004
Male	1.764	1.298	2.399	<0.001		
Female	1					
Age by gender (male)	0.988	0.981	0.995	0.001	0.001	<0.001
Ethnicity					0.032	0.105
White	1					
Black	2.823	2.438	3.268	<0.001		
Asian	1.671	1.215	2.298	0.002		
Mixed	1.839	1.276	2.626	0.001		
Other	1.504	1.210	1.869	<0.001		
Index diagnosis					0.127	0.196
ARMS (c)	1					
Acute and transient psychotic disorders	2.682	1.981	3.631	<0.001		
Substance use disorders	0.146	0.105	0.202	<0.001		
Bipolar mood disorders	0.839	0.598	1.178	0.310		
Non bipolar mood disorders	0.152	0.109	0.210	<0.001		
Anxiety disorders	0.107	0.077	0.148	<0.001		
Personality disorders	0.213	0.141	0.321	<0.001		
Developmental disorders	0.031	0.015	0.064	<0.001		
Childhood/adolescence onset disorders	0.039	0.025	0.061	<0.001		
Physiological syndromes	0.085	0.052	0.137	<0.001		
Mental retardation	0.086	0.049	0.151	<0.001		

(a) The C-index of the overall model was 0.800 (95% CI 0.785-0.816);

(b) a base model was used which included only the predictor age; the C-index for the base model was 0.567 (95%CI 0.552-0.581);

(c) when the APS subgroup of the ARMS was used as reference group the overall results were unchanged (likelihood ratio chi-square test=1772.83, p<0.001):

°Age: 1.012 (95%CI 1.007 - 1.017);

°Gender: Female 1, Male 1.775 (95%CI 1.305 - 2.415);

°Age by Gender: Male .988 (95%CI .980 - .995);

°Ethnicity: White 1, Black 2.819 (95%CI 2.436 - 3.264), Asian 1.683 (95%CI 1.223 - 2.314), Mixed 1.827 (95%CI 1.274 - 2.621), Other 1.504 (95%CI 1.210 - 1.869);

°Index diagnosis: APS 1, BLIPS 1.944 (95%CI 1.053 - 3.590), GRD not estimated (5 cases and 0 events), Acute and transient psychotic disorders 3.054 (95% 2.162 - 4.315), Substance use disorders 0.166 (95%CI 0.115 - 0.239), Bipolar mood disorders 0.956 (95%CI 0.655 - 1.396), Non bipolar mood disorders 0.173 (95%CI 0.119 - 0.249), Anxiety disorders 0.122 (95%CI 0.085 - 0.176), Personality disorders 0.242 (95%CI 0.156 - 0.377), Developmental disorders 0.035 (95%CI 0.017 - 0.074), Childhood/adolescence onset disorders 0.044 (95%CI 0.027 - 0.071), Physiological syndromes 0.097 (95%CI 0.059 - 0.161), Mental retardation 0.098 (95%CI 0.055 - 0.177).

Table 3. Performance of the risk calculator for transdiagnostic prediction of psychosis in secondary mental health care and clinical usefulness

Performance measure	Derivation			Validation		
Overall						
Brier (a)	0.025			0.019		
R^2 (mean, 95% CI)	0.747	0.705	0.786	0.72	0.673	0.762
Discrimination						
Harrell's C (mean, 95%CI)	0.800	0.785	0.816	0.791	0.778	0.807
Discrimination slope (mean, 95% CI)	1.465	1.412	1.518	1.408	1.353	1.462
Calibration						
Calibration-in-the-large	0			0.005		
Calibration slope (mean, 95%CI)	1			0.956	0.909	1.003
Clinical usefulness						
Net benefit at 7.69%(b)	0.053			0.061		

a) at 10-year, b) compared to treat all at 5-year

Figure 2. Decision curve analysis estimated in the validation dataset, showing the potential clinical usefulness of the risk calculator at different threshold probabilities (risk of developing psychosis by 5 years) for focused interventions to prevent psychosis, as compared to treating all patients or to treating no patients at all.